

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

- (i) APPLICANT: Hillman, Jennifer L.  
Lal, Preeti  
Corley, Neil C.  
Shah, Purvi
- (ii) TITLE OF THE INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL  
(4,5) BISPHOSPHATE 5-PHOSPHATASE
- (iii) NUMBER OF SEQUENCES: 5
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Incyte Genomics, Inc.  
(B) STREET: 3160 Porter Drive  
(C) CITY: Palo Alto  
(D) STATE: CA  
(E) COUNTRY: USA  
(F) ZIP: 94304
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: To Be Assigned  
(B) FILING DATE: Filed Herewith
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: 09/258,643; 08/884,681  
(B) FILING DATE: February 26, 1999; June 27, 1997
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Eckstrom, Richard A.  
(B) REGISTRATION NUMBER: 37,027  
(C) REFERENCE/DOCKET NUMBER: PF-0334-2 DIV
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 650-855-0555  
(B) TELEFAX: 650-845-4166

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRSTNOT03  
(B) CLONE: 638789

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Met Asp Val Leu Ser Pro Leu Ser Phe Ile Lys Val Ser His Val Arg
 1             5             10             15
Met Gln Gly Ile Leu Leu Leu Val Phe Ala Lys Tyr Gln His Leu Pro
          20          25          30
Tyr Ile Gln Ile Leu Ser Thr Lys Ser Thr Pro Thr Gly Leu Phe Gly
          35          40          45
Tyr Trp Gly Asn Lys Gly Gly Val Asn Ile Cys Leu Lys Leu Tyr Gly

```

50 55 60  
Tyr Tyr Val Ser Ile Ile Asn Cys His Leu Pro Pro His Ile Ser Asn  
65 70 75 80  
Asn Tyr Gln Arg Leu Glu His Phe Asp Arg Ile Leu Glu Met Gln Asn  
85 90 95  
Cys Glu Gly Arg Asp Ile Pro Asn Ile Leu Asp His Asp Leu Ile Ile  
100 105 110  
Trp Phe Gly Asp Met Asn Phe Arg Ile Glu Asp Phe Gly Leu His Phe  
115 120 125  
Val Arg Glu Ser Ile Lys Asn Arg Cys Tyr Gly Gly Leu Trp Glu Lys  
130 135 140  
Asp Gln Leu Ser Ile Ala Lys Lys His Asp Pro Leu Leu Arg Glu Phe  
145 150 155 160  
Gln Glu Gly Arg Leu Leu Phe Pro Pro Thr Tyr Lys Phe Asp Arg Asn  
165 170 175  
Ser Asn Asp Tyr Asp Thr Ser Glu Lys Lys Arg Lys Pro Ala Trp Thr  
180 185 190  
Asp Arg Ile Leu Trp Arg Leu Lys Arg Gln Pro Cys Ala Gly Pro Asp  
195 200 205  
Thr Pro Ile Pro Pro Ala Ser His Phe Ser Leu Ser Leu Arg Gly Tyr  
210 215 220  
Ser Ser His Met Thr Tyr Gly Ile Ser Asp His Lys Pro Val Ser Gly  
225 230 235 240  
Thr Phe Asp Leu Glu Leu Lys Pro Leu Val Ser Ala Pro Leu Ile Val  
245 250 255  
Leu Met Pro Glu Asp Leu Trp Thr Val Glu Asn Asp Met Met Val Ser  
260 265 270  
Tyr Ser Ser Thr Ser Asp Phe Pro Ser Ser Pro Trp Asp Trp Ile Gly  
275 280 285  
Leu Tyr Lys Val Gly Leu Arg Asp Val Asn Asp Tyr Val Ser Tyr Ala  
290 295 300  
Trp Val Gly Asp Ser Lys Val Ser Cys Ser Asp Asn Leu Asn Gln Val  
305 310 315 320  
Tyr Ile Asp Ile Ser Asn Ile Pro Thr Thr Glu Asp Glu Phe Leu Leu  
325 330 335  
Cys Tyr Tyr Ser Asn Ser Leu Arg Ser Val Val Gly Ile Ser Arg Pro  
340 345 350  
Phe Gln Ile Pro Pro Gly Ser Leu Arg Glu Asp Pro Leu Gly Glu Ala  
355 360 365  
Gln Pro Gln Ile  
370

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2573 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: BRSTNOT03
- (B) CLONE: 638789

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAAGGCTCAG	CATACACGTC	GTGACTTGGA	ACGTGGCTTC	GGCAGCGCCC	CTCGAGCTCT	60
CAGTGACCTG	CTTCAGCTGA	ACAACCGGAA	CCTCAATCTT	GACATATATG	TTATTGGTTT	120
GCAGGAATTG	AACTCTGGGA	TCATAAGCCT	CCTTTCCGAT	GCTGCCTTTA	ATGACTCGTG	180
GAGCAGTTTC	CTCATGGATG	TGCTTTCCCC	TCTGAGCTTC	ATCAAGGTCT	CCCATGTCCG	240
TATGCAGGGG	ATCCTCTTAC	TGGTCTTTGC	CAAGTATCAG	CATTTGCCCT	ATATCCAGAT	300
TCTGTCTACT	AAATCCACCC	CCACTGGCCT	GTTTGGGTAC	TGGGGGAACA	AAGGTGGAGT	360
CAACATCTGC	CTGAAGCTTT	ATGGCTACTA	TGTCAGCATC	ATCAACTGCC	ACCTGCCTCC	420
CCACATTTCC	AACAATTACC	AGCGGCTGGA	GCACTTTGAC	CGGATCCTGG	AGATGCAGAA	480
TTGTGAGGGG	CGAGACATCC	CAAACATCCT	GGACCACGAC	CTCATTATCT	GGTTTGGAGA	540
CATGAACCTT	CGGATCGAGG	ACTTTGGGTT	GCACTTTGTT	CGGGAATCCA	TTAAAAATCG	600

GTGCTACGGT GGCCTGTGGG AGAAGGACCA GCTCAGCATT GCCAAGAAAC ATGACCCGCT 660  
 GCTCCGGGAG TTCCAGGAGG GCCGCCTACT CTTCCCGCCC ACCTACAAGT TTGATAGGAA 720  
 CTCCAACGAC TATGACACCA GTGAGAAAAA ACGCAAGCCT GCATGGACCG ATCGCATCCT 780  
 GTGGAGGCTG AAGCGGCAGC CCTGTGCTGG CCCCAGACT CCCATACCGC CGGCCTCACA 840  
 CTTCTCCTTG TCTCTGAGGG GCTACAGCAG CCACATGACG TACGGCATCA GCGACCACAA 900  
 GCCTGTCTCC GGCACGTTCC ACTTGGAGCT GAAGCCATTG GTGTCTGCTC CGCTGATCGT 960  
 CCTGATGCCC GAGGACCTGT GGACCGTGGA AAATGACATG ATGGTCAGCT ACTCTTCAAC 1020  
 CTCGGACTTC CCCAGCAGCC CGTGGGACTG GATTGGACTG TACAAGGTGG GGCTGCGGGA 1080  
 CGTTAATGAC TACGTGTCTT ATGCCTGGGT CCGGGACAGC AAGGTCTCCT GCAGCGACAA 1140  
 CCTGAACCAG GTTTACATCG ACATCAGCAA TATCCCTACC ACTGAAGATG AGTTTCTCCT 1200  
 CTGTTACTAC AGCAACAGTC TCGGTTCTGT GGTGGGGATA AGCAGACCCT TCCAGATCCC 1260  
 GCCTGGCTCC TTGAGGGAGG ACCCACTGGG TGAAGCACAG CCACAGATCT GAGCCAGGAT 1320  
 GGGAGTGAAT CCCAGGCGGA GGCCAGAGCT GGCAGCCAGC TCTGCCTTTC CACTGCCGGG 1380  
 AGTGCTGGGG GCCCAGCCTG GCCCCCTGAA GAGACAGCCA AGTGTCTGTC ACATACTCCT 1440  
 CCCAGAGTGA GCTCTAACCA GGCTCATTTG CTCTCTCCAC TACTCATCTC TGGAAATTAGC 1500  
 CGCTTAAATA CAGGTTTTTG TTGCTGAGAT GTGAGTGAAA CCAGCTAGTG TGTCAACAGT 1560  
 GAAGACCTGG GGACAGTTCT GCGTCTCATT TCTGGATTCC TACCCCTCTCT TCTAGTCTTG 1620  
 CCCAAGTAGT CCTGCCAGGC ACATGCCCCA TTTGGCACAG GCCTGCATTCT TTGTCTGTCC 1680  
 GTCCTGGGCC TCAGGCTGTC TGGGAGGGGA GATGCTCACA TTTGTACAGG CTACATAGAC 1740  
 TGGTGCAAGC AGTGCTGGAT TCCAGGAGTC TTGGCATCTC ATAGCTTGTC CCCGTGAGGA 1800  
 GTGAGCAGAG GGTCTGGGAT TTCTGCTTTT AGCAAAAGCA GTCTGACTCA GTGGGCAGAA 1860  
 TGGAGGGGCC CCTCTAGCCA GGCTCTTACG CCATGGTTAT GAGCAGGTTG ATGAGGGTCC 1920  
 TTCGGCCAGC ACAACCTTCC TCCCTACTCA CGGCATGGAG TCTGACTGCA TGGAAAGTCC 1980  
 AGATCCTGAC AGAGAGAACT GGAAGGATC CAGGTTTCGT TCCGTTGGTA GCTTGAGTCC 2040  
 CATGCCTCCA CCCTGCCATC TGAGGAAGGG GTGACAAGTG GTCAAGGAGC TGTGGCCACA 2100  
 GACTTTTCCA GGGTGGTCCT TGCCAGGTGA GGTGCGTCTG TGCCACCCTT GTCAGGAGTC 2160  
 ATTGACGACG GCGCCCCCTT GGACCCCCCG GGACCTCAGA GTGGGGGCAG GCAGAAGGGA 2220  
 GAACCAGCTC AAGACATTTT GGAGGATCTG GCCCTGGGGT TCTTCAGAGA ACACCCTCTA 2280  
 GGGGCTTTGG GGACATGGCC TGTCCCCACA TCCAGCACTT GCCTCCGCCA TGGTCACTCG 2340  
 GCAGCCCTTT TCCCAGGAGA AGACACCTCT GGGAGCCTGC TCAGTGCTTG TCCTGCCATC 2400  
 CTGTGTCTG GGAAGTGGG TTAAGTCCAGT TGCTCTGTGT TGCATACTCT CCCCCGCAAG 2460  
 CCTGTGTATG AAGAATTGTC CCCTGGCTTC CAGCAGGCCA TGGCTGGCTG TTTTGTGACT 2520  
 GTTACATTGT GCAGGGGTAA TTATTAGCGT GGCTTTTACA CTAAAAAAA AAA 2573

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank  
 (B) CLONE: 1399101

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala Gly Leu Phe Gly Tyr Trp Gly Asn Lys Gly Gly Val Asn Ile Cys  
 1 5 10 15  
 Leu Lys Leu Tyr Gly Tyr Tyr Val Ser Ile Ile Asn Cys His Leu Pro  
 20 25 30  
 Pro His Ile Ser Asn Asn Tyr Gln Arg Leu Glu His Phe Asp Arg Ile  
 35 40 45  
 Leu Glu Met Gln Asn Cys Glu Gly Arg Asp Ile Pro Asn Ile Leu Asp  
 50 55 60  
 His Asp Leu Ile Ile Trp Phe Gly Asp Met Asn Phe Arg Ile Glu Asp  
 65 70 75 80  
 Phe Gly Leu His Phe Val Arg Glu Ser Ile Lys Asn Arg Cys Tyr Gly  
 85 90 95  
 Gly Leu Trp Glu Lys Asp Gln Leu Ser Ile Ala Lys Lys His Asp Pro  
 100 105 110  
 Leu Leu Arg Glu Phe Gln Glu Gly Arg Leu Leu Phe Pro Thr Tyr  
 115 120 125  
 Lys Phe Asp Arg Asn Ser Asn Asp Tyr Asp Thr Ser Glu Lys Lys Arg

130	135	140
Lys Pro Ala Trp Thr Asp Arg Ile Leu Trp Arg Leu Lys Arg Gln Pro		
145	150	155
Cys Ala Gly Pro Asp Thr Pro Ile Pro Pro Ala Ser His Phe Ser Leu		160
	165	170
Ser Leu Arg Gly Tyr Ser Ser His Met Thr Tyr Gly Ile Ser Asp His		175
	180	185
Lys Pro Val Ser Gly Thr Phe Asp Leu Glu Leu Lys Pro Leu Val Ser		190
	195	200
Ala Pro Leu Ile Val Leu Met Pro Glu Asp Leu Trp Thr Val Glu Asn		205
	210	215
Asp Met Met Val Ser Tyr Ser Ser Thr Ser Asp Phe Pro Ser Ser Pro		220
225	230	235
Trp Asp Trp Ile Gly Leu Tyr Lys Val Gly Leu Arg Asp Val Asn Asp		240
	245	250
Tyr Val Ser Tyr Ala Trp Val Gly Asp Ser Lys Val Ser Cys Ser Asp		255
	260	265
Asn Leu Asn Gln Val Tyr Ile Asp Ile Ser Asn Ile Pro Thr Thr Glu		270
	275	280
Asp Glu Phe Leu Leu Cys Tyr Tyr Arg Asn Ser Leu Arg Ser Val Val		285
	290	295
Gly Ile Arg Arg Pro Phe Gln Ile Pro Pro Gly Ser Leu Arg Glu Asp		300
305	310	315
Pro Leu Gly Glu Ala Gln Pro Gln Ile		320
	325	

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 942 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1019103

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Val Thr Val Pro Glu Pro Gly Ala Ala Glu Ser Arg Ala Pro Cys Gly	
1 5 10 15	
Asp Ser Ser Gly Cys Val Arg Ser Ala Gly Ala Ser Met Asp Gln	
20 25 30	
Ser Val Ala Ile Gln Glu Thr Leu Ala Glu Gly Glu Tyr Cys Val Ile	
35 40 45	
Ala Val Gln Gly Val Leu Cys Glu Gly Asp Ser Arg Gln Ser Arg Leu	
50 55 60	
Leu Gly Leu Val Arg Tyr Arg Leu Glu His Gly Gly Gln Glu His Ala	
65 70 75 80	
Leu Phe Leu Tyr Thr His Arg Arg Met Ala Ile Thr Gly Asp Asp Val	
85 90 95	
Ser Leu Asp Gln Ile Val Pro Val Ser Arg Asp Phe Thr Leu Glu Glu	
100 105 110	
Val Ser Pro Asp Gly Glu Leu Tyr Ile Leu Gly Ser Asp Val Thr Val	
115 120 125	
Gln Leu Asp Thr Ala Glu Leu Ser Leu Val Phe Gln Leu Pro Phe Gly	
130 135 140	
Ser Gln Thr Arg Met Phe Leu His Glu Val Ala Arg Ala Cys Pro Gly	
145 150 155 160	
Phe Asp Ser Ala Thr Arg Asp Pro Glu Phe Leu Trp Leu Ser Arg Tyr	
165 170 175	
Arg Cys Ala Glu Leu Glu Leu Glu Met Pro Thr Pro Arg Gly Cys Asn	
180 185 190	
Ser Ala Leu Val Thr Trp Pro Gly Tyr Ala Thr Ile Gly Gly Gly Gly	
195 200 205	

Ser	Asn	Phe	Asp	Gly	Leu	Arg	Pro	Asn	Gly	Lys	Gly	Val	Pro	Met	Asp
210						215					220				
Gln	Ser	Ser	Arg	Gly	Gln	Asp	Lys	Pro	Glu	Ser	Leu	Gln	Pro	Arg	Gln
225					230					235					240
Asn	Lys	Ser	Lys	Ser	Glu	Ile	Thr	Asp	Met	Val	Arg	Ser	Ser	Thr	Ile
				245					250					255	
Thr	Val	Ser	Asp	Lys	Ala	His	Ile	Leu	Ser	Met	Gln	Lys	Phe	Gly	Leu
			260					265					270		
Arg	Asp	Thr	Ile	Val	Lys	Ser	His	Leu	Gln	Lys	Glu	Glu	Asp	Tyr	
		275					280				285				
Thr	Tyr	Ile	Gln	Asn	Phe	Arg	Phe	Phe	Ala	Gly	Thr	Tyr	Asn	Val	Asn
		290				295					300				
Gly	Gln	Ser	Pro	Lys	Glu	Cys	Leu	Arg	Leu	Trp	Leu	Ser	Asn	Gly	Ile
305					310					315					320
Gln	Ala	Pro	Asp	Val	Tyr	Cys	Val	Gly	Phe	Gln	Glu	Leu	Asp	Leu	Ser
				325					330					335	
Lys	Glu	Ala	Phe	Phe	Phe	His	Asp	Thr	Pro	Lys	Glu	Glu	Glu	Trp	Phe
			340					345					350		
Lys	Ala	Val	Ser	Glu	Gly	Leu	His	Pro	Asp	Ala	Lys	Tyr	Ala	Lys	Val
		355					360					365			
Lys	Leu	Ile	Arg	Leu	Val	Gly	Ile	Met	Leu	Leu	Leu	Tyr	Val	Lys	Gln
		370				375					380				
Glu	His	Ala	Ala	Tyr	Ile	Ser	Glu	Val	Glu	Ala	Glu	Thr	Val	Gly	Thr
385					390					395					400
Gly	Ile	Met	Gly	Arg	Met	Gly	Asn	Lys	Gly	Gly	Val	Ala	Ile	Arg	Phe
				405					410					415	
Gln	Phe	His	Asn	Thr	Ser	Ile	Cys	Val	Val	Asn	Ser	His	Leu	Ala	Ala
			420					425					430		
His	Ile	Glu	Glu	Tyr	Glu	Arg	Arg	Asn	Gln	Asp	Tyr	Lys	Asp	Ile	Cys
		435					440					445			
Ser	Arg	Met	Gln	Phe	Cys	Gln	Pro	Asp	Pro	Ser	Leu	Pro	Pro	Leu	Thr
		450				455					460				
Ile	Ser	Asn	His	Asp	Val	Ile	Leu	Trp	Leu	Gly	Asp	Leu	Asn	Tyr	Arg
465				470						475					480
Ile	Glu	Glu	Leu	Asp	Val	Glu	Lys	Val	Lys	Lys	Leu	Ile	Glu	Glu	Lys
				485					490					495	
Asp	Phe	Gln	Met	Leu	Tyr	Ala	Tyr	Asp	Gln	Leu	Lys	Ile	Gln	Val	Ala
			500					505					510		
Ala	Lys	Thr	Val	Phe	Glu	Gly	Phe	Thr	Glu	Gly	Glu	Leu	Thr	Phe	Gln
			515					520				525			
Pro	Thr	Tyr	Lys	Tyr	Asp	Thr	Gly	Ser	Asp	Asp	Trp	Asp	Thr	Ser	Glu
		530				535					540				
Lys	Cys	Arg	Ala	Pro	Ala	Trp	Cys	Asp	Arg	Ile	Leu	Trp	Lys	Gly	Lys
545					550					555					560
Asn	Ile	Thr	Gln	Leu	Ser	Tyr	Gln	Ser	His	Met	Ala	Leu	Lys	Thr	Ser
				565					570					575	
Asp	His	Lys	Pro	Val	Ser	Ser	Val	Phe	Asp	Ile	Gly	Val	Arg	Val	Val
			580					585				590			
Asn	Asp	Glu	Leu	Tyr	Arg	Lys	Thr	Leu	Glu	Glu	Ile	Val	Arg	Ser	Leu
			595				600					605			
Asp	Lys	Met	Glu	Asn	Ala	Asn	Ile	Pro	Ser	Val	Ser	Leu	Ser	Lys	Arg
			610			615					620				
Glu	Phe	Cys	Phe	Gln	Asn	Val	Lys	Tyr	Met	Gln	Leu	Lys	Val	Glu	Ser
625					630					635					640
Phe	Thr	Ile	His	Asn	Gly	Gln	Val	Pro	Cys	His	Phe	Glu	Phe	Ile	Asn
				645					650					655	
Lys	Pro	Asp	Glu	Glu	Ser	Tyr	Cys	Lys	Gln	Trp	Leu	Asn	Ala	Asn	Pro
			660					665					670		
Ser	Arg	Gly	Phe	Leu	Leu	Pro	Asp	Ser	Asp	Val	Glu	Ile	Asp	Leu	Glu
			675				680					685			
Leu	Phe	Val	Asn	Lys	Thr	Thr	Ala	Thr	Lys	Leu	Asn	Ser	Gly	Glu	Asp
			690			695					700				
Lys	Ile	Glu	Asp	Ile	Leu	Val	Leu	His	Leu	Asp	Arg	Gly	Lys	Asp	Tyr
705					710					715					720
Phe	Leu	Ser	Val	Ser	Gly	Asn	Tyr	Leu	Pro	Ser	Cys	Phe	Gly	Ser	Pro
				725					730					735	

Ile His Thr Leu Cys Tyr Met Arg Glu Pro Ile Leu Asp Leu Pro Leu  
 740 745 750  
 Glu Thr Ile Ser Glu Leu Thr Leu Met Pro Val Trp Thr Gly Asp Asp  
 755 760 765  
 Gly Ser Gln Leu Asp Ser Pro Met Glu Ile Pro Lys Glu Leu Trp Met  
 770 775 780  
 Met Val Asp Tyr Leu Tyr Arg Asn Ala Val Gln Gln Glu Asp Leu Phe  
 785 790 795 800  
 Gln Gln Pro Gly Leu Arg Ser Glu Phe Glu His Ile Arg Asp Cys Leu  
 805 810 815  
 Asp Thr Gly Met Ile Asp Asn Leu Ser Ala Ser Asn His Ser Val Ala  
 820 825 830  
 Glu Ala Leu Leu Leu Phe Leu Glu Ser Leu Pro Glu Pro Val Ile Cys  
 835 840 845  
 Tyr Ser Thr Tyr His Asn Cys Leu Glu Cys Ser Gly Asn Tyr Thr Ala  
 850 855 860  
 Ser Lys Gln Val Ile Ser Thr Leu Pro Ile Phe His Lys Asn Val Phe  
 865 870 875 880  
 His Tyr Leu Met Ala Phe Leu Arg Glu Leu Lys Asn Ser Ala Lys  
 885 890 895  
 Asn His Leu Asp Glu Asn Ile Leu Ala Ser Ile Phe Gly Ser Leu Leu  
 900 905 910  
 Leu Arg Asn Pro Ala Gly His Gln Lys Leu Asp Met Thr Glu Lys Lys  
 915 920 925  
 Lys Ala Gln Glu Phe Ile His Gln Phe Leu Cys Asn Pro Leu  
 930 935 940

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 901 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 1420920

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Glu Pro Pro Leu Pro Val Gly Ala Gln Pro Leu Ala Thr Val Glu  
 1 5 10 15  
 Gly Met Glu Met Lys Gly Pro Leu Arg Glu Pro Cys Ala Leu Thr Leu  
 20 25 30  
 Ala Gln Arg Asn Gly Gln Tyr Glu Leu Ile Ile Gln Leu His Glu Lys  
 35 40 45  
 Glu Gln His Val Gln Asp Ile Ile Pro Ile Asn Ser His Phe Arg Cys  
 50 55 60  
 Val Gln Glu Ala Glu Glu Thr Leu Leu Ile Asp Ile Ala Ser Asn Ser  
 65 70 75 80  
 Gly Cys Lys Ile Arg Val Gln Gly Asp Trp Ile Arg Glu Arg Arg Phe  
 85 90 95  
 Glu Ile Pro Asp Glu Glu His Cys Leu Lys Phe Leu Ser Ala Val Leu  
 100 105 110  
 Ala Ala Gln Lys Ala Gln Ser Gln Leu Leu Val Pro Glu Gln Lys Asp  
 115 120 125  
 Ser Ser Ser Trp Tyr Gln Lys Leu Asp Thr Lys Asp Lys Pro Ser Val  
 130 135 140  
 Phe Ser Gly Leu Leu Gly Phe Glu Asp Asn Phe Ser Ser Met Asn Leu  
 145 150 155 160  
 Asp Lys Lys Ile Asn Ser Gln Asn Gln Pro Thr Gly Ile His Arg Glu  
 165 170 175  
 Pro Pro Pro Pro Phe Ser Val Asn Lys Met Leu Pro Arg Glu Lys  
 180 185 190  
 Glu Ala Ser Asn Lys Glu Gln Pro Lys Val Thr Asn Thr Met Arg Lys

58

Val	Pro	Leu	Asp	Glu	Gly	Ala	Ser	Glu	Arg	Pro	Leu	Gln	Val	Pro	Lys
				725					730					735	
Glu	Ile	Trp	Leu	Leu	Val	Asp	His	Leu	Phe	Lys	Tyr	Ala	Cys	His	Gln
			740					745					750		
Glu	Asp	Leu	Phe	Gln	Thr	Pro	Gly	Met	Gln	Glu	Glu	Leu	Gln	Gln	Ile
		755					760					765			
Ile	Asp	Cys	Leu	Asp	Thr	Ser	Ile	Pro	Glu	Thr	Ile	Pro	Gly	Ser	Asn
	770					775					780				
His	Ser	Val	Ala	Glu	Ala	Leu	Leu	Ile	Phe	Leu	Glu	Ala	Leu	Pro	Glu
785				790					795						800
Pro	Val	Ile	Cys	Tyr	Glu	Leu	Tyr	Gln	Arg	Cys	Leu	Asp	Ser	Ala	Tyr
			805					810						815	
Asp	Pro	Arg	Ile	Cys	Arg	Gln	Val	Ile	Ser	Gln	Leu	Pro	Arg	Cys	His
		820					825					830			
Arg	Asn	Val	Phe	Arg	Tyr	Leu	Met	Ala	Phe	Leu	Arg	Glu	Leu	Leu	Lys
	835					840					845				
Phe	Ser	Glu	Tyr	Asn	Ser	Val	Asn	Ala	Asn	Met	Ile	Ala	Thr	Leu	Phe
	850					855				860					
Thr	Ser	Leu	Leu	Leu	Arg	Pro	Pro	Pro	Asn	Leu	Met	Ala	Arg	Gln	Thr
865					870				875						880
Pro	Ser	Asp	Arg	Gln	Arg	Ala	Ile	Gln	Phe	Leu	Leu	Gly	Phe	Leu	Leu
			885					890						895	
Gly	Ser	Glu	Glu	Asp											
			900												

1092267 10360